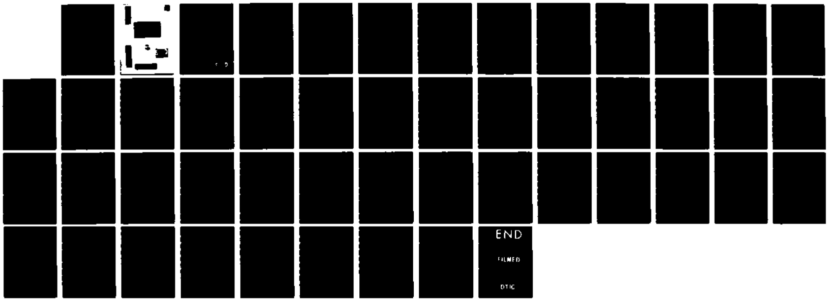
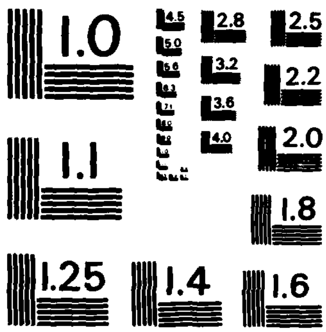
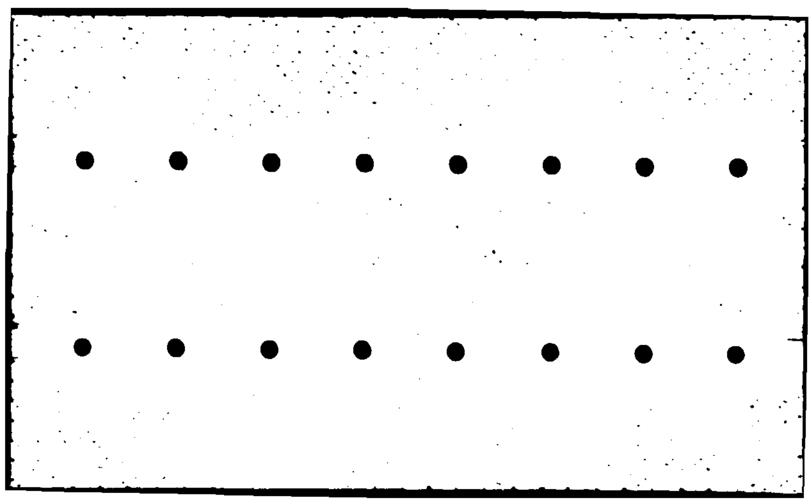
AD-A160 634   WHY BOTHER WITH EXPERIMENTS?(U) CHICAGO UNIV IL CENTER   1/1
FOR DECISION RESEARCH R M HOGARTH MAY 85 TR-16
N00014-84-C-0018

UNCLASSIFIED   F/G 14/4   NL





MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS-1963-A



12

WHY BOTHER WITH EXPERIMENTS?

Robin M. Hogarth
Graduate School of Business
University of Chicago
Center for Decision Research

May 1985

DRAFT--COMMENTS WELCOME

Sponsored by:
Engineering Psychology Programs
Office of Naval Research
Contract Number N00014-84-C-0018
Work Unit Number NR 197-080

Unclassified
SECURITY CLASSIFICATION OF THIS PAGE

REPORT DOCUMENTATION PAGE

1a. REPORT SECURITY CLASSIFICATION Unclassified	1b. RESTRICTIVE MARKINGS
2a. SECURITY CLASSIFICATION AUTHORITY	3. DISTRIBUTION / AVAILABILITY OF REPORT Approved for public release; distribution unlimited.
2b. DECLASSIFICATION / DOWNGRADING SCHEDULE	
4. PERFORMING ORGANIZATION REPORT NUMBER(S) 16	5. MONITORING ORGANIZATION REPORT NUMBER(S)
6a. NAME OF PERFORMING ORGANIZATION Center for Decision Research Graduate School of Business University of Chicago 6b. OFFICE SYMBOL (If applicable)	7a. NAME OF MONITORING ORGANIZATION Engineering Psychology Programs
6c. ADDRESS (City, State, and ZIP Code) 1101 East 58th Street Chicago, IL 60637	7b. ADDRESS (City, State, and ZIP Code) Code #442EP Arlington, VA 22217
8a. NAME OF FUNDING / SPONSORING ORGANIZATION Office of Naval Research 8b. OFFICE SYMBOL (If applicable)	9. PROCUREMENT INSTRUMENT IDENTIFICATION NUMBER N00014-84-C-0018
8c. ADDRESS (City, State, and ZIP Code) Department of the Navy Arlington, VA 22217	10. SOURCE OF FUNDING NUMBERS PROGRAM ELEMENT NO. / PROJECT NO. / TASK NO. / WORK UNIT ACCESSION NO. NR-197-080

11. TITLE (Include Security Classification)
Why Bother with Experiments?

12. PERSONAL AUTHOR(S)
Robin M. Hogarth

13a. TYPE OF REPORT	13b. TIME COVERED	14. DATE OF REPORT (Year, Month, Day)	15. PAGE COUNT
Technical Report	FROM ______ TO ______	May 1985	43

16. SUPPLEMENTARY NOTATION

17. COSATI CODES			18. SUBJECT TERMS (Continue on reverse if necessary and identify by block number)
FIELD	GROUP	SUB-GROUP	Generalization; experimental evidence; model building.

19. ABSTRACT (Continue on reverse if necessary and identify by block number)

A generalization is a working hypothesis, typically expressed in the form of cause-effect relations. In the social sciences, generalizations decay because (a) it is difficult to identify appropriate cause-effect relations, and (b) such relations are sensitive to the influences of environmental conditions. Whereas scientists should be realistic in their aspirations to create generalizable knowledge, much can be done to improve performance through the use of formal models and experimentation. It is particularly important that theories permit comparisons between models and data at multiple levels involving processes, environmental conditions, and predictions. Scientists should avoid the extremes of "models without data" and "data without models." Instead, models whould be subjected to "strong" empirical tests via predictions (rather than tests of statistical significance), and the competing predictions of alternatives. In addition to suggesting what experimental evidence should be collected, models also serve the important function of determining when data collection would be of little value. The nature of experimental evidence is considered from

20. DISTRIBUTION / AVAILABILITY OF ABSTRACT ☐ UNCLASSIFIED/UNLIMITED ☒ SAME AS RPT. ☐ DTIC USERS	21. ABSTRACT SECURITY CLASSIFICATION Unclassified	
22a. NAME OF RESPONSIBLE INDIVIDUAL	22b. TELEPHONE (Include Area Code)	22c. OFFICE SYMBOL

DD FORM 1473, 84 MAR — 83 APR edition may be used until exhausted. All other editions are obsolete.

SECURITY CLASSIFICATION OF THIS PAGE
Unclassified

19. ABSTRACT

three viewpoints: (1) asymmetries in the way data and models interact in affecting conclusions; (2) apparent but illusory conflicts between the goals of internal and external validity; and (3) the importance of conducting experiments despite poor prospects of creating knowledge that can be generalized.

Why Bother with Experiments?

1. Introduction

Since the participants at this conference are scientists, I would like to share with you the conversation I once had with one of my daughters, then 10 years old:

"Daddy, scientists discover things, don't they?

"Yes."

"Daddy, are you a scientist?"

"Yes."

"Well, what have you discovered?"

". . . ."

My inability to answer this question was a considerable blow to my self-esteem until I realized that most social scientists would have been left equally speechless. More importantly, it made me reflect on the difficulties of generating knowledge in the social sciences, the methods we use, and the ephemeral nature of our conclusions.

The purpose of this paper is to elaborate on these thoughts; it is organized as follows. In section 2, I consider the difficulties involved in creating knowledge that can be generalized. This involves asking what is meant by the term "generalization" and why generalizations are so short-lived in the social sciences. The creation of knowledge is a painstaking enterprise and, whereas humility in aspirations should be the rule, much can still be done to increase the efficiency of scientific endeavors. Ways of gaining knowledge through the use of formal models and experimentation are discussed from this viewpoint in Sections 3 and 4, respectively. To anticipate the sequel, I argue that models and data must interact at all phases of scientific

investigation. Too many efforts fall into what I call the categories of "models without data" (e.g., parts of modern economics) and "data without models" (e.g., much of social psychology). I also answer the question posed in the title of this paper, i.e., Why bother with experiments? Specifically, I do not advocate any particular type of experiment but believe in the utility of multiple methods of data collection going from mathematical simulations to artificial laboratory tasks to quite complex field studies. I also advocate multiple methods. In my view, the appropriate experimental approach depends in large part on both the nature of the phenomenon being studied and the state of theory or model development. Moreover, I shall elaborate several reasons why I believe we should bother about experiments. Throughout this paper, I shall support my arguments with examples of research that are known to me, primarily through my interests in the psychology of judgment and decision making. This inevitably leads to a parochial view on these issues for which I ask the reader's indulgence. On the other hand, in research it is difficult to discuss the <u>how</u> without considering the <u>what</u>.

2. Generalizations decay

In a provocative paper, Cronbach (1975) wrote:

> Generalizations decay. At one time a conclusion describes the existing situation well, at a later time it accounts for rather little variance, and ultimately it is only valid as history (pp. 122-123).

The decision making literature is full of generalizations at various stages of decay. For example, "people are risk averse," "people ignore base rates", "there is a confirmation bias in hypothesis testing," or "people prefer non-ambiguous probabilities in choice." Cronbach's words raise three critical questions: (1) What is meant by generalization? (2) Why do generalizations decay? and (3) What can we do about this situation?

(1) One way of clarifying the meaning of generalization is to step back and raise the issue of the purpose of scientific investigation. Perhaps simplistically, I see science as the creation of knowledge which is most usefully codified in terms of causal statements. These cause-effect relations may be elaborated in more or less detail and can be stated in deterministic or probabilistic terms. Implicit in our search to make these statements is the belief that the complexity of nature is capable of explanation in relatively simple terms. The appeal of simplicity (or parsimony) of explanation is twofold: first, simple explanations invoke a sense of wonder when they account for complex phenomena; and second, simple explanations are easy to remember and use. In my view, the two hallmarks of good science are beauty and utility.

The cause-effect relations advanced by scientists in the form of "generalizations" have usually evolved through a cyclical process that involves (a) observation of effects, (b) speculation as to the causes of effects (otherwise known as generating hypotheses or building models), and (c) further observation (possibly including experimentation) leading to further speculation, and so on. The important point about this process is that generalizations made at any particular moment are nothing more than working hypotheses (see also Cronbach, 1975, p. 125). Some working hypotheses do, of course, work better and last longer than others. However, it is essential to bear in mind that our generalizations (however dearly cherished) are nothing more than working hypotheses.

(2) Why do generalizations (working hypotheses) decay? I like to think about this in the following way. Statements of cause and effect are useful to the extent that they bring order into our understanding of the world. Such order, however, is achieved at the cost of simplification. Since it is

impractical to have theories that are totally realistic, we are forced to "satisfice" (cf. Simon, 1979). This implies that although we like to think of the world as being governed by simple cause-effect relations of the type illustrated in Figure 1 (simple generalization), it is more accurately described by relations exhibited in Figure 2. Note from Figure 2 that, in the

Insert Figures 1, 2, and 3 about here

real world, simple cause-effect relations only hold when certain conditions are present or absent, i.e., when causes are conjoined by specific conditions. For example, will striking a match produce a flame? Yes, but only in the presence of oxygen.

In developing hypotheses, our major inferential problem is that we typically first notice effects and then have to reason backwards to try and infer underlying cause(s). However, to the extent that underlying cause-effect relations are modified by environmental conditions, our ability to make these inferences is complicated. Indeed, there is often considerable ambiguity concerning whether and when particular variables are causes or conditions, or perhaps both (cf. Mackie, 1974).

How do these ideas apply to understanding the psychology of decision making? First, note that this essentially involves explaining how relatively simple organisms (i.e., humans) manage to cope with infinitely more complex environments. Thus, if like myself, you believe that people draw upon a limited number of strategies or principles for making decisions (admittedly often in complex combinations), the inferential problem typically faced by researchers is that depicted in Figure 3. From this framework, it is easy to see both why it is difficult to infer cause-effect relations and why generalizations decay. First, whereas effects are typically observable, the underlying cause(s) may be unobservable. Also it is not evident that

Figure 1

$$\text{Cause} \longrightarrow \text{Effect}$$

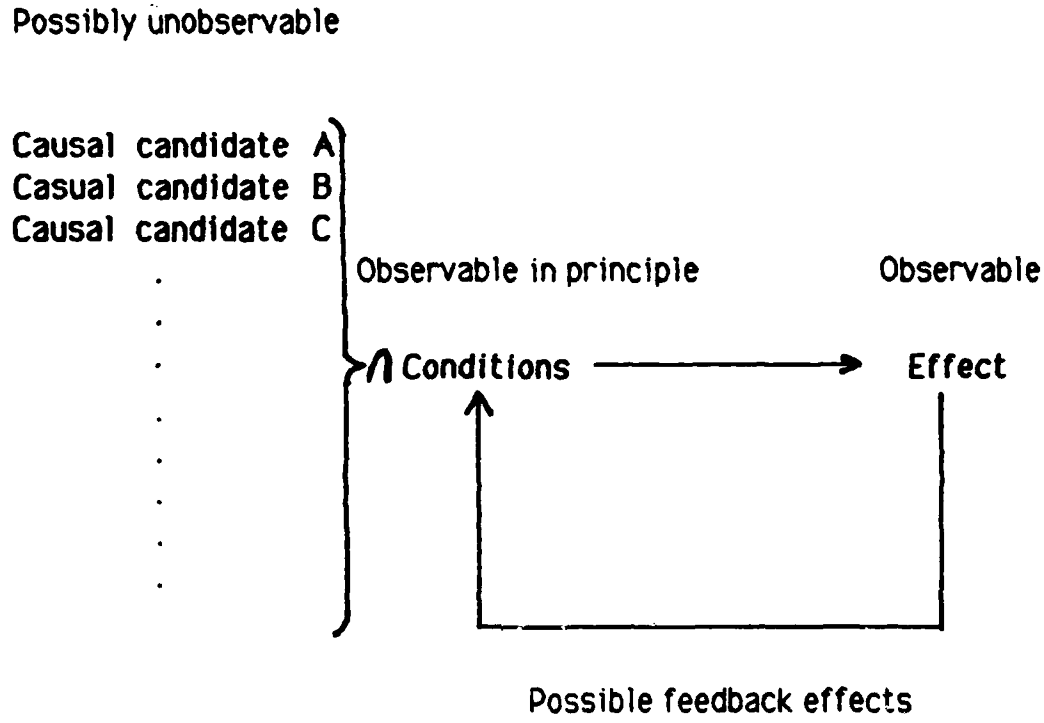
How we like to think of the world -- Simple generalization

Figure 2

$$\text{Cause} \wedge \text{Conditions} \longrightarrow \text{Effect}$$

A more accurate statement of the world -- Conditional generalization

Figure 3



Structure of inferential problem typically faced by researchers

researchers will infer the appropriate causal agent (or behavioral principle) from observing effects. Second, even if the appropriate principle is inferred, the influences of environmental conditions need to be assessed. Moreover, whereas different conditions are typically observable in principle, in practice it may require many variations in experimental/observational circumstances before one can determine the relative importance of different conditions on effects. Third, effects can--via feedback loops--sometimes influence the conditions in which they occur thereby both changing the importance of the latter and even influencing the likelihood of their own occurrence (Bandura, 1978; Maruyama, 1963). I now briefly consider these points.

Identifying causal agents. The process by which people identify causal agents, and thus build hypotheses or models, is a major topic about which I can only comment briefly here. (However, see Einhorn & Hogarth, 1982a; 1985; Hogarth, 1982). Leaving aside the tricky epistemological questions of what does or does not constitute a "cause," consider three of the major complexities of this process. First, there is the sheer physical difficulty of being able to select one or several variables from the mass of potentially available information. It is my opinion that not many scientists work at this level and thus it is easy to forget (with hindsight) just how difficult it is to do this successfully. For example, how did Pasteur come to the then totally foreign realization that invisible microbes cause disease, the effects of which can be both very visible and dramatic? One answer is that most discoveries are informed by prior theories, however loosely specified. But this misses the point of how these theories evolved in the first place. The second difficulty relates precisely to the nature of the theories used to direct the search for variables. What if these are misguided? Consider, for example, the false

predictions implied by Newtonian physics, nineteenth century notions of blood circulation, diet, and so on. The third difficulty results from the second. This occurs when we take actions based on theories that are false and these actions, in turn, prevent us from learning that our theories are false (cf. Einhorn & Hogarth, 1978). Lewis Thomas (1983) provides an example of such a theory being applied in an "operational context." The proponent was a notable physician at the beginning of this century:

> This physician enjoyed the reputation of a diagnostician, with a particular skill in diagnosing typhoid fever, then the commonest disease on the wards of New York's hospitals. He placed particular reliance on the appearance of the tongue, which was universal in the medicine of that day (now entirely inexplicable, long forgotten). He believed that he could detect significant differences by palpating that organ. The ward rounds conducted by this man were, essentially, tongue rounds; each patient would stick out his tongue while the eminence took it between thumb and forefinger, feeling its texture and irregularities, then moving from bed to bed, diagnosing typhoid in its earliest stages over and over again, and turning out a week or so later to have been right, to everyone's amazement. He was a more productive carrier, using only his hands, than Typhoid Mary (p. 22).

Lest my comments on this topic appear unduly pessimistic, let me also refer you to Campbell (1960). In a fascinating article on the creative process, he points out that much problem solving activity inevitably involves a painful process of trial and error. As he states:

> The tremendous number of non-productive thought trials . . . must not be underestimated. Think of what a small proportion of thought becomes conscious, and of conscious thought what a small proportion gets uttered, what a still smaller fragment gets published, and what a small proportion of what is published is used by the next intellectual generation. There is a tremendous wastefulness, slowness, and rarity of achievement (Campbell, 1960, p. 393).

Environmental conditions. Figure 3 illustrates the importance of environmental conditions on the range of applicability of general statements. That is, since observed effects result from the interaction of processes (causal agents) and environmental conditions, it is necessary both

to specify and understand the influences of the latter. Cronbach (1975) also makes the point that our ability to make enduring generalizations depends heavily on the nature of the environment confronted by our simple notions of cause and effect. The notion used by Cronbach is that of the difference between closed and open systems. Thus:

> The half-life of an empirical proposition may be great or small. The more open a system, the shorter the half-life of relations within it are likely to be . . . . Propositions describing atoms and electrons have a long half-life, and the physical theorist can regard the processes in his world as steady. Rarely is a social or behavioral phenomenon isolated enough to have this steady-process property. Hence the explanations we live by will perhaps always remain partial and distant from real events . . . . and rather short lived (Cronbach, 1975, p. 123).

In behavioral decision making, the most dramatic results in recent years have demonstrated precisely how sensitive subjects' responses are to seemingly minor changes in tasks (Einhorn & Hogarth, 1981). This, in turn, has led to greater appreciation of the task-contingent nature of strategies for judgment and choice (Payne, 1982) and has important implications for research. First, since complex behavior at the overt level is not necessarily inconsistent with simple underlying processes, this suggests seeking a limited number of theoretical principles to explain the underlying or covert responses that initiate behavior. Parenthetically I have often thought that it would be particularly interesting if these psychological principles were found to have physiological counterparts. For example, Coombs and Avrunin's (1977) notions of "goods satiating" and "bads escalating" seem to capture the way we process both psychological and physical pleasures and pains.

Second, to understand this behavior, it becomes critical to vary environmental responses in systematic ways and to analyze carefully the task conditions in which the behavior is observed. As Ken Hammond has already stated, one needs to sample behavior across both persons and situations and to

respect the ranges that variables normally take in the environment (see also Hammond, 1978). Cronbach (1975) also stresses the role of task description, going so far as to state:

> Instead of making generalization the ruling consideration in our research, I suggest that we reverse our priorities. An observer collecting data in one particular situation is in a position to appraise a practice or proposition in that setting, observing effects in context. In trying to describe and account for what happened, he will give attention to whatever variables were controlled, but he will give equally careful attention to uncontrolled conditions, to personal characteristics, and to events that occurred during treatment and measurement. As he goes from situation to situation, his first task is to describe and interpret the effect anew in each locale, perhaps taking into account factors unique to that local series of events. . . . That is, generalization comes late, and the exception is taken as seriously as the rule. (Cronbach, 1975, pp. 124-125.)

I disagree with Cronbach in so far as I believe that attempts at generalization (or forming working hypotheses) are useful even at early stages of research. However, his emphasis on attempting to understand variations in task conditions is exemplary. Indeed, the importance of these ideas can be illustrated by noting that "theories" often have an unfortunate tendency to "asymptote" on explaining surprising phenomena that have been generated within fairly tight environmental circumstances. For example, whereas I greatly admire Kahneman and Tversky's (1979) prospect theory, I recognize that it cannot handle certain phenomena that occur when you do something as simple as change the size of payoffs. (Specifically, prospect theory predicts certain violations of expected utility theory when payoffs are small. However, I know of two unpublished studies where this prediction fails). Nonetheless, this sub-field of choice theory is currently populated with many models that seek to explain the same data reported by Kahneman and Tversky. I hate, for instance, to think of the number of published, let alone unpublished explanations of the Allais paradox. The theoretical challenge is not simply

to explain the current "anomalies"; rather it is to make predictions that take us beyond these phenomena. However, this can only be done if investigators construct their models by considering implications for a wider range of environmental circumstances than has been the case to date.

(3) To summarize, generalizations are working hypotheses expressed in terms of cause-effect relations. Generalizations decay because (a) it is difficult to identify appropriate causal agents (working hypotheses) and (b) observations of simple cause-effect relations are complicated by the myriad of environmental conditions in which these operate. Given what has been stated, it is legitimate to question whether we are or ever will be equipped in the social sciences to produce generalizations capable of resisting decay. Science, however, involves both costs and benefits. Thus whereas we should be more realistic concerning individual aspirations and costs, we should also recall that the benefits of science do not lie in discovery. The benefits lie in application. It takes only one or a few people to discover something once; but a discovery can be used on countless occasions. This is not to say that all is well in the way social science is conducted. Indeed, by careful attention to methods, much can be done to prevent generalizations from decay as well as accelerating this process when necessary. I therefore now consider the principal means used to generate knowledge. These are, respectively, the development of formal models and the use of experimentation.

3. The role of formal models

Formal models have a dual role; first, to extend our understanding of working hypotheses; and second, to delimit the extent of our knowledge and, in some cases, even to show where it is impossible to acquire knowledge.

A model is a concise statement of a working hypothesis. It can be, but is not necessarily expressed in mathematical form. Good models have two

characteristics: (1) Economy of description; and (2) The power to suggest implications that are not evident at first sight. In other words, a model is like an intellectual crutch. It enables the scientist to go further. Whereas models are abstractions, the critical dimension by which we evaluate models is their ability to generate insight about naturally occurring phenomena. That is models have to relate to data.

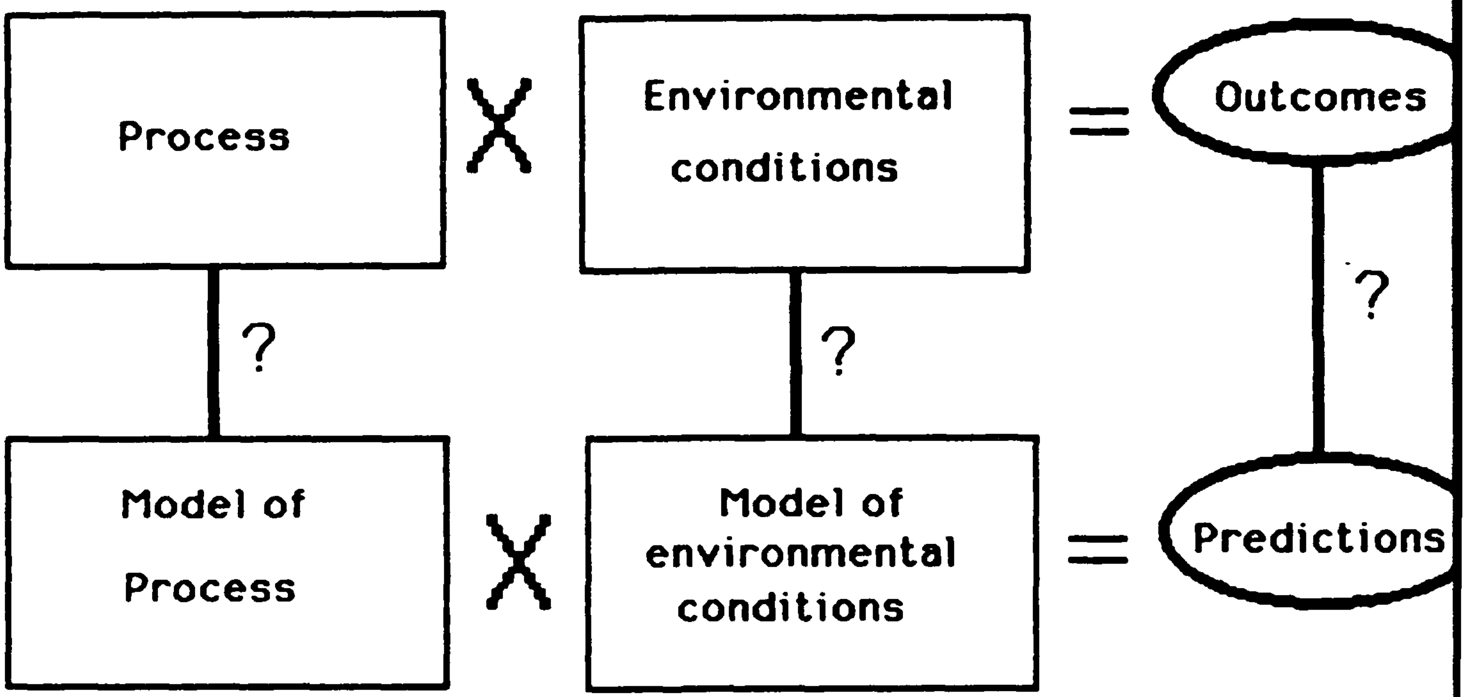
It is important to realize that models can relate to data at various levels. To illustrate, consider the "lens model"-like (Brunswik, 1952) diagram in Figure 4. This elaborates on Figure 3 in that it suggests that

Insert Figure 4 about here

for any "true" process (represented by data of differing types) the scientist can build models that permit comparisons between models and data at several different levels. Roughly speaking, these comparison points are at (a) the level of the assumed underlying process, (b) concerning environmental conditions, and (c) predictions versus observations. In my view, the better models in the social sciences permit comparisons at all three levels such that one can make several types of judgment concerning model validity. To appreciate this, consider the lack of progress in areas where attempts are not made to create these links. There are two extreme cases: "models without data" (a good part of modern economics, cf. Kuttner, 1985) and "data without models" (a good part of social psychology).

Models without data. Data, or observations, are usually at the origin of models. That is, based on observations the scientist makes assumptions about the underlying causal agent presumed to generate the phenomena. However, this is not always done. In economics, for example, working hypotheses about underlying processes are frequently invoked in the form of "as if" assumptions without any regard for known facts, i.e., data. Whereas this is often

Figure 4



Points of correspondence between data (process) and model

practical from a modeling viewpoint, I question whether such models could ever generalize (see also below) unless the "as if" assumptions can also be shown to simulate what is known about the facts at that level. For example, it is evident that the utility maximizing assumptions made on behalf of _homo economicus_ do not square with what we know about limitations on human information processing abilities. The troublesome aspect of this, however, is not so much this lack of correspondence as such, but the failure to show that the "as if" assumptions might imply some correspondence. (For an example of how "as if" linear models of judgment can simulate more complex underlying processes, see Einhorn, Kleinmuntz and Kleinmuntz, 1979).

Second, given the importance of environmental conditions, useful data-model comparisons can also be made at this level. Once again, one may choose to ignore this potential source of reality testing by declaring certain conditions to be irrelevant, as in the treatment of contextual variables in expected utility theory (von Neumann & Morgenstern, 1947). However, given the importance of environmental conditions, the relative success of many models in the social sciences is determined by the way in which environmental conditions are represented, irrespective of assumptions about process (see also below).

Third, data or observations can inform models at the level of predictive accuracy. Many scientists, myself included, put great weight on the criterion of predictive accuracy. Some, for example Milton Friedman (1956), argue that the predictive accuracy of a model is all that matters. That is, one need not worry whether working hypotheses are "realistic" provided one's model "yields predictions that are good enough for the purpose at hand or that are better than predictions from alternative theories" (Friedman, 1956, p. 41). Unfortunately, generalizations decay and, of late, it cannot be said that the economic models espoused by Friedman have escaped this fate. Moreover, as can

be seen by consulting Figure 4, it is unclear what one should do when prediction fails if one's models do not also permit comparisons between data and model at the levels of both process and environmental conditions. (For a possible precedent concerning what people do "when prophecy fails," see Festinger, Riecken & Schachter, 1956.)

Data without models. On considering this extreme, I am reminded of a quote from Pirandello: "A fact is like a sack. It won't stand up unless you put something in it." In other words, the reporting of data or results always assumes some underlying theoretical notions. Indeed, a result can only be surprising if it violates expectations and thus one's theories about the world (cf. Davis, 1971). However, even if facts do have some surprise value, the meaning (and thus potential for generalization) of such facts is not evident unless they are accompanied by some specific underlying model or theory. For example, whereas the robustness of the "conjunction fallacy" reported by Tversky and Kahneman (1983) is a fascinating finding, it is not clear what this tells us about human cognitive processes except that these can produce outcomes that are inconsistent with the prescriptions of probability theory. At this level, of course, the conjunction fallacy is hardly a new finding. Note that Tversky and Kahneman are not saying that human judgment always shows conjunction effects. (Indeed, this would be a simple "generalization.") On the other hand, by failing to develop a model, they are unable to inform us both of how these judgments are made and of the conditions under which people do or do not commit this "fallacy." As stated by Abelson (1984), "You can really only say that you understand a phenomenon when you can make it go away." However, to "make it go away" you need a model.

Formal models can contribute to different aspects of the scientific process. Consider their potential roles in (1) observing data, (2) specifying

implications prior to data collection, and (3) simulation.

(1) Whereas formal models are usually associated with fairly advanced states of scientific investigation, as noted above there are also important links to be made between data and models at early stages. I would like to emphasize that the observation of data requires considerable theoretical skill, the need for which has been delightfully captured by Louis Guttman (1982) when talking of "exploratory" research. As Guttman stated, one does not send novices to the North Pole or the moon since they would not know what to look for. For example, it took a Darwin to develop evolutionary theory even though people had observed the multitude of animal species for centuries before him. Similarly, whereas fragments of ancient bones may mean little to you or me, they could have profound significance for paleontologists. Closer to my own interests, verbal protocols of subjects involved in problem solving or decision making tasks do not mean much unless you know what you are looking for. Protocols do not automatically generate theories.

(2) After initial observation, formal models are important to the gathering of new evidence. The most critical task here is to suggest predictions or implications for empirical testing. This is particularly interesting when implications violate intuition (Davis, 1971).

Testing may be done in either a "weak" or a "strong" sense (cf. Platt, 1964). The weak form occurs when investigators try to assess whether data are or are not consistent with a particular hypothesis. For example, does the behavior of stock prices conform to one of the criteria of market efficiency (cf. Fama, 1970)? This form of testing models is weak in the sense that it usually involves some variation of the "null hypothesis trap." That is, tests center on whether or not the data disconfirm the null hypothesis. (For a recent discussion, see Serlin & Lapsley, 1985.) Thus, when studies use

naturally occurring field data, it is often the case that measurement and other sources of noise prohibit clean tests of the null hypothesis. For example, in discussing with a leading "rational" economist whether the expected utility model could ever be disconfirmed by naturally occuring data (involving real decision makers facing real payoffs), he admitted that in practice this would be extremely difficult. However, since in principle it was not impossible, the model was "falsifiable" and thus defensible as such. In psychological experiments, on the other hand, the data are typically cleaner. However, a simple "acceptance" or rejection of the null hypothesis at some conventional level of statistical significance usually gives little idea of the substantive importance of the model's predictions.

Strong tests depend critically on developing models and have two characteristics. First, models are judged by predictions that avoid entrapment by the null hypothesis. Note that these predictions need not take the form of precise numbers; they could be qualitative in nature, involve the specification of approximate functional forms, and so on. Indeed, in discussing the "slow progress of soft psychology," Meehl (1978) decries the reliance on statistical hypothesis testing, advocating instead the following "moral":

> It is always valuable to show approximate agreement of observations with a theoretically predicted numerical point value, rank, order, or function form, than it is to compute a "precise probability" that something merely differs from something else (Meehl, 1978, p. 825, emphasis omitted).

Second, strong tests also require the specification of plausible alternative models and the delineation of conditions where the alternatives make different predictions. In the absence of "plausible" alternatives, even the use of naive baseline models would improve practice considerably. In the social sciences, however, these practices are the exception rather than the rule. For example, in a survey of empirical papers published in <u>Management Science</u>

from 1955 to 1976, Armstrong (1979) found that less than one quarter (22%) considered more than one hypothesis.

A good example of the use of alternative models is provided by Thaler and Shefrin (1983). They challenged conventional economic notions of savings behavior based both on Friedman's permanent income hypothesis and Modigliani's life cycle model by positing a more psychologically plausible hypothesis involving notions of self-control. Their paper is essentially a review. Specifically, they test the predictions of the alternative models on the data reported in various studies published in the literature. In my view, the Thaler-Shefrin model provides a better account of the data than the more conventional "working hypotheses." However, this is not the point. The point is that it does not suffice to say that conventional theories are implausible and even to show that they do not fit the facts; it is necessary to build alternatives. For example, many people (myself included) admire Howard Kunreuther's fine field study of insurance decision making (Kunreuther et al., 1978). In a nutshell, Kunreuther found that people's decisions concerning the purchase of flood and earthquake insurance were inconsistent with the expected utility model. However, apart from appealing after the fact to concepts such as availability (Tversky & Kahneman, 1973), Kunreuther did not formulate a precise alternative model of the insurance purchasing decision that could have been rejected by the data collected. (On the other hand, it is true that Kunreuther's study provides data that can inform the building of alternative models. See, e.g., Hogarth & Kunreuther, 1985.)

To summarize, there are at least three advantages to developing specific alternative models: (a) The yield from experiments is greatly enhanced if the data provide information about more than one working hypothesis; (b) As noted with respect to Guttman's comment, theoretical skills require expertise. How-

ever, expertise does not develop in a vacuum. Developing specific models requires practice in developing models and thus I see this emphasis as potentially increasing the level of scientific reasoning; (c) Whereas generalizations decay, old theories never die unless they are replaced. That is, the most powerful way of disconfirming a hypothesis is to replace it with one that predicts better. Developing specific alternatives is essential to the process of regeneration.

(3) Whereas I have emphasized the need for interaction between data and models, there are areas in which models can make important contributions and yet only be loosely connected to data. These all involve various forms of "simulation."

One type of simulation is the theoretical exploration of working hypotheses aimed at exploring conditions under which these do or do not hold. A recent and instructive example is provided by Klayman and Ha (1985). Klayman and Ha took on one of the most unquestioned and apparently robust generalizations in the decision making literature. This is the so-called "confirmation bias" (Wason & Johnson-Laird, 1972) whereby, when testing hypotheses, people are said to have a deleterious tendency to seek information that could confirm but not disconfirm their beliefs. By careful theoretical modeling of this task, Klayman and Ha show conditions under which a confirmation strategy is, in fact, the more appropriate approach. The surprising finding is that these conditions are relatively common. Thus, in one theoretical paper, and by asking a question about conditions, Klayman and Ha illuminate what had become a whole research tradition in experimental psychology. It is important to note, incidentally, that the Klayman-Ha paper does not model the process by which humans test hypotheses. Rather, it shows the consequences of using various strategies in different environments. As

such, it is paradoxically more illuminating about the psychology of hypothesis testing than many studies that attempt to describe "what people do." On the other hand, the nature of this paradox is resolved when one considers the importance of conditions on behavior as illustrated in Figure 3. Sometimes generalization decay can be usefully accelerated by good theoretical analysis.

A second type of simulation occurs when an investigator uses models (computer or mathematical) to mimic the behavior of people in order to investigate the implications of behavioral assumptions in different environments. A good example of this type of work is that of Axelrod (1984) on the evolution of cooperation. Axelrod investigated the survival rates of different competitive strategies involved in repeated plays of the prisoners' dilemma game. His observations consisted of pitting different strategies against each other and noting which did more or less well, on average, against all opponents. From observing the relative performance of different types of strategies, he inferred characteristics that were more or less likely to foster survival in different types of environments. From my view, whereas Axelrod collected no "real" data, this was a particularly illuminating study in that it suggests hypotheses for considering more complex, natural situations where controlled experimentation would be difficult, if not infeasible, to conduct. For example, the purchasing arms of many corporations are essentially involved in repeated prisoners' dilemma games with their suppliers. Axelrod's methodology and results, although not conclusive, are rich in suggestions for the implications of adopt -, different strategies. It is also significant that Axelrod's work has subsequently inspired work in the natural sciences investigating the survival strategies of birds (Lombardo, 1985).

In my view, one of the most important ideas in modern science is the notion of the impossibility theorem introduced by Gödel. This is also a form of simulation, the full power of which, however, is insufficiently appreciated. The classic example in the social sciences is Arrow's (1963) result concerning the aggregation of preferences. Arrow showed that there is no way to aggregate individual preferences (meeting certain specifications) in a manner that does not violate at least one desirable postulate of rationality. However, this kind of reasoning could be taken much further. For example, reconsider Figure 3 and imagine that you are trying to model a particular phenomenon. In many cases it should be possible to show that the complexity of the environment in which the phenomenon occurs is such that the power of any generalization must be very weak. To be able to do this, however, requires developing a model of the underlying phenomenon. That is, without a model one cannot assess whether any working hypothesis has the slightest chance of surviving a change in environmental conditions. As an example of this strategy, I was once interested in investigating claims that a certain proportion of variance in IQ could be attributed to heredity (Hogarth, 1974). Rather than question the data, I chose to examine the underlying model by considering its assumptions. What I found (via simulation) was that minor changes in some questionable assumptions had huge effects on estimates of variance components. In other words, even if you could collect good data, it was unclear what you could infer from them. To summarize, I see the "impossibility" approach to simulation as extremely revealing and powerful concerning when and what data to collect. Once again, it requires the ability (and willingness) to specify models. Unfortunately, it is considerably underutilized.

4. Why we should bother about experiments

Above I have discussed the importance of formal models. On considering how experimental evidence can affect our ability to generalize, I now consider (1) how models and data interact in affecting conclusions, (2) apparent conflicts between the goals of internal and external validity, and (3) reasons why we should persist in doing experiments even though prospects for achieving generality are typically poor.

(1) Whereas the tradition of resolving issues by critical experiments is well established in physical science, this is not the case in the social sciences. Indeed, many are quite skeptical about what can be achieved in this area via experimentation. The reasons, I believe, relate to the issues discussed in the previous section. These are, (a) the relative lack of well specified models and alternatives, and (b) the conditions under which experiments are conducted raise serious issues as to how results can be generalized. Moreover, in conducting and interpreting experiments, these issues are not unrelated.

I have already discussed the need for well specified models and alternatives. However, as an example of how the presence or absence of such models interacts with experimental methodology, consider the following two experiments. One involved extremely weak methodology, but no one doubts the effect. The other study used a sophisticated design, controls, and so on, and although there were effects, people still remain skeptical as to their generality. The first study took place near Los Alamos, New Mexico, in July 1945 and involved a huge explosion. The study itself was "weak" by social science standards. There was but a single observation and no control group. However, no-one doubts that the explosion was caused by detonating an atomic bomb. The second study was much more sophisticated. This was the so-called New Jersey

"negative income tax" experiment (Kershaw, 1975). In this carefully designed study, an attempt was made to determine the effect on the labor supply of a negative income tax program for the poor. The study involved paying negative taxes to participating families. There were randomly selected experimental and control groups. Subjects were enrolled in different locations, and so on. However, despite all methodological precautions, it is still possible to be skeptical about the claims made concerning the possible effects of instituting a negative income tax on a larger scale. Indeed, as we all know, faith in the outcome of the first experiment led to operational implementation. However, this can not be said of the second. There are many differences between the experiments in New Mexico and New Jersey. The first dealt with a physical phenomenon for which theories existed and where prior experience had been codified (even though considerable uncertainty existed prior to the experiment about the size of the probable effect); there was high agreement on what variables were relevant; and credible alternative explanations to account for the large explosion are hard to imagine (e.g., the possibility of a huge subterranean earthquake occurring at precisely that time and place seems remote). The second study, on the other hand, took place in conditions where prior theory was far looser and rival interpretations were almost bound to plague the results of even the best of research designs.

(2) In attempts to generalize theories, researchers often have to grapple with issues of how "realistic" or "representative" experiments need to be. This too leads to interesting interactions with the degree of commitment people have toward particular theories. Since Campbell and Stanley's (1966) classic work, the issue of realism in experiments has often been conceptualized as one of internal versus external validity. One can considerably reduce (if not eliminate) threats to internal validity by careful controls. However,

in doing so one may preclude the possibility of reaching conclusions that are externally valid, i.e., the objectives of internal and external validity trade off. In discussing the nature of this trade-off, Swieringa and Weick (1982) have argued that, contrary to most researchers' beliefs, the real indifference function between internal and external validity may be convex rather than concave. Thus researchers who add a little realism to artificial laboratory experiments can diminish rather than increase the utility of their efforts. To support this view, they cite the work of Plott and others in experimental economics (see e.g., Plott, 1982; Smith, 1982). Here the tradition has been to test economic theory in the most abstract conditions possible. The argument is that the principles of economic theory are abstract, i.e., "content free." They should therefore be tested in abstract environments since, if they do not hold here, they are surely suspect in more complex real world situations. At first sight, this argument has much appeal. However, consider the reactions one is liable to elicit if such experiments produce results that alternatively (a) confirm or (b) disconfirm the principles tested. Results that confirm the theoretical status quo will undoubtedly find ready acceptance. Indeed, in these cases the abstract "artificiality" of the experiments is seen as a plus ("a nice, clean test"). On the other hand, disconfirming results can be easily dismissed as irrelevant. That is, artificiality has now become a negative ("the study was not representative"). For example, although there has been considerable interest in demonstrating that rats and pigeons obey the laws of supply and demand (Battalio, Green & Kagel, 1981), would the same interest exist if experiments had shown that they didn't? Indeed, would we even be aware that these experiments had been performed? Artificiality is often advantageous; however, it can lead to considerable selectivity in what ultimately becomes public knowledge. (This

is not meant to imply that selectivity is otherwise absent in science. For specific evidence, see Mahoney, 1977.) Parenthetically, I have often wondered whether, in addition to sharing respect for the laws of supply and demand, rats and pigeons are also subject to preference reversals and other biases. If they were, what would we make of these findings?

In many ways, the dilemma between internal and external validity is false. To see why, reconsider Figure 3. As noted above, the goal of science is to make statements of cause and effect and to understand how these are modified by environmental conditions. Thus, to understand the nature of some phenomenon, it is important to vary environmental conditions. This therefore implies conducting a range of "experiments" going from highly controlled laboratory conditions to quite "loose" field studies. Thus, I do not see conflict between experimental approaches that focus on internal as opposed to external validity, or vice versa, <u>provided both approaches are part of the same scientific program</u>. On the other hand, what is unforgivable is to form strong beliefs in working hypotheses that have only been tested in one type of environment. Unfortunately, the way science is organized (and in particular the way scientists are rewarded), individual scientists tend to focus on the implications of their own work within limited environments, and to seek outlets for publication within their own disciplines in journals that reward narrow empirical approaches.

In my view, we lack studies where scientists from different backgrounds have attempted to examine the same phenomena or working hypotheses in different environments. Fortunately, some exceptions "confirm this rule." Consider, for example, the approach taken by Kunreuther et al. (1978) in studying insurance decision making. This involved both economists doing field studies and psychologists doing laboratory experiments. Alternatively,

consider some of Thaler's (1980) interesting field observations that were inspired by Kahneman and Tversky's (1979) prospect theory. There is little doubt that if tentative cause-effect relations prove to be robust across different environmental conditions, they are granted greater respect (cf. Einhorn & Hogarth, 1985). In medicine, for example, it is standard practice to verify whether hypotheses based on observations in one population also hold in others. Note that I am not saying that the same social scientists should be involved in doing studies across ranges of all possible conditions. On the other hand, I advocate exploring means whereby we encourage the testing of hypotheses across areas and thus across different conditions.

Summarizing, I have argued that the way to test the "generalization" of working hypotheses is by testing across different conditions. That is, I advocate multiple forms of experimentation dealing with the same topic. By the same token, I also advocate multiple methods. For example, I am tired of debates concerning the utility of "process tracing" data. It is healthy that different researchers should study similar topics by different research techniques. What is unhealthy is that any group should consider its approach as intrinsically more realistic than those of others. As a good example of multiple methods in the area of decision making, consider the work of Einhorn, Kleinmuntz and Kleinmuntz (1979). They showed that protocol methods <u>and</u> analytic models can be profitably used to study cognitive processes since they illuminate different aspects and levels of the same phenomena. It is my hope that in the future researchers within different methodological traditions can learn that truth can (perhaps?) be shared (cf. Einhorn & Hogarth, 1981).

(3) As noted above, prospects for achieving generalizations that resist decay are remote for most of the social sciences. Indeed, in some areas people question whether one should even do experiments. Despite the problems,

let me offer six reasons why we should continue to do experiments:

(i) A little knowledge is better than none. Whereas one should be realistic concerning what can be achieved via experimentation, some partial knowledge and half-truths can be gained. For example, experiments conducted in the late 1960s and early 1970s pointed to glaring deficiencies in human ability to process probabilistic information (e.g., Tversky & Kahneman, 1974). Whereas these results led to extreme over-generalizations in some quarters (e.g., Nisbett & Ross, 1980), recent years have seen a growing realization of the conditions under which people do or do not make certain inferential errors (e.g., Nisbett et al., 1983) or whether some response tendencies are indeed "biased" (e.g., Hogarth, 1981). To my mind, the critical factor is not that we do or do not state working hypotheses so much as we know what weight to accord to them. We must educate ourselves not to expect too much of experiments but bear in mind that the possession of knowledge is relative. As stated by Erasmus, "In the land of the blind, the one-eyed man is king."

(ii) Experiments help avoid metaphysical speculation. Two contributory causes of the explosion of knowledge in the physical and natural sciences in the past two centuries have been (a) the growth of methods for conducting experiments and (b) the physical means of doing so. For example, whereas the notion of "a control group" is almost second nature to today's practising scientists, it is significant that this critical idea is of fairly recent origin (Boring, 1954). In addition, one tends to forget how the development of computational equipment has, in addition to enabling scientists even to consider certain issues, facilitated the penetration of statistical methods developed only in this century. Without experimentation, or even possibilities for experimentation, scientists are not afforded the important

posssibility of learning that their ideas (sometimes called "theories") are typically wrong. Although, as noted above, I favor greater rigor in the formulation of models than is generally the case today, I am equally vehement about stating that science cannot depend too heavily on axiomatic reasoning since, by definition, this only applies within restricted, closed worlds that lack the open systems characteristics of our everyday reality. Tautological reasoning, that is the basis of all implications of logical systems and therefore highly useful, is limited. In particular, there is no guarantee that tautological truths are empirically valid (Einhorn & Hogarth, 1982b).

(iii) Experiments can and should be used to illuminate scientific conflicts. Whereas the topic of scientific dispute has received much attention in recent years, with some scientists even advocating adversarial methods (so-called "science courts"), in my view we are all better served by defining critical experiments. To illustrate, assume that a conflict exists between the proponents of two rival theories. If the theories differ, then they must make different predictions concerning events that are yet to be observed. The test, therefore, is to require the rivals to define these events, make predictions, and then collect the appropriate experimental evidence. The critical aspect of this process is that, <u>prior</u> to conducting the experiment, the opponents must agree both on what evidence should be collected, and how different possible results should be interpreted. For further elaboration, and more radical propositions concerning these ideas, see Hofstee (1984).

(iv) Experiments can impact on practice. Restricting my comments to the decision making literature, many experiments can be and are usefully conducted without specifically attempting to establish behavioral laws. Consider, for example, the problems of assessing judgmental inputs for management science

models. Much practical knowledge can be established by determining which methods subjects seem to prefer in making the required judgments. This is not to say that I am advocating totally atheoretical approaches to data collection. However, if data were routinely collected within an experimental framework, this would undoubtedly help the subsequent development of theory. In particular, I believe this would increase sensitivity to the effects of the conditions under which the data are collected.

(v) Experiments are a form of history. In one of the quotations from Cronbach (1975) given above, the concept of the half-life of a finding was used. However, even it some of today's truths have a short half-life, documenting their existence can be extraordinarily important. Each generation views it reality through eyes trained by its predecessors. Thus, in understanding the issues and perspectives of today, it is essential to trace how such matters developed over time in order to understand why they are deemed important. For somebody trying to understand how decision researchers (or any other group) view certain issues at a given time, the existence of an experimental literature is of enormous importance. Experiments are one way in which a science expresses its values and concerns and also attempts to describe its empirical reality. Once again, I quote Cronbach (1975):

> The special task of the social scientist in each generation is to pin down the contemporary facts. Beyond that, he shares with the humanistic scholar and the artist in the effort to gain insight into contemporary relationships, and to realign the culture's views of man with present realities. To know man as he is is no mean aspiration (p. 126).

Parenthetically, I emphasize that I use the word "experiments" in a broad sense to include field observations and case studies as well as the narrower laboratory tasks favored by psychologists. Although often dismissed by "rigorous methodologists," case studies can be most informative; however, their yield depends crucially upon whether the investigator adopts an

"experimental framework" in organizing observations (for further discussion on this point, see Campbell, 1975).

(vi) Experiments help define new questions. Given the complexity of empirical phenomena, it is rare that experiments provide definitive "answers." Indeed, it is an old adage that good research raises more questions than it answers. Experiments can thus be a vital source of good questions. The importance of this should not be underestimated. It was pertinently raised some time ago by Gertrude Stein when she mused: "Suppose no one asked a question. What would the answer be?" More interestingly, Einstein stressed the importance of <u>doing</u> when engaged in inquiry and made the following distinction between a scientist and a detective: "For the detective, the crime is given, the problem posed: Who killed Cock Robin? The scientist must at least in part commit his own crime." (Einstein & Infeld, 1938, p. 76).

References

Abelson, R. P. (1984). Remark made during colloquium at Department of Psychology, Northwestern University.

Armstrong, J. S. (1979). Advocacy and objectivity in science. _Management Science_, _25_, 423-428.

Arrow, K. J. (1963). _Social choice and individual values_. (2nd edition). New Haven: Yale University Press.

Axelrod, R. (1984). _The evolution of cooperation_. New York: Basic Books.

Bandura, A. (1978). The self system in reciprocal determinism. _American Psychologist_, _33_, 344-358.

Battalio, R., Green, L., & Kagel, J. (1981). Income-leisure tradeoffs of animal workers. _American Economic Review_, _71_, 621-632.

Boring, E. G. (1954). The nature and history of experimental control. _American Journal of Psychology_, _67_, 573-589.

Brunswik, E. (1952). _The conceptual framework of psychology_. Chicago: University of Chicago Press.

Campbell, D. T. (1960). Blind variation and selective retention in creative thought as in other knowledge processes. _Psychological Review_, _67_, 380-400.

Campbell, D. T. (1975). "Degrees of freedom" and the case study. _Comparative Political Studies_, _8_, 178-193.

Campbell, D. T., & Stanley, J. C. (1966). _Experimental and quasi-experimental designs for research_. Chicago: Rand-McNally.

Coombs, C. H., & Avrunin, G. (1977). Single-peaked functions and the theory of preference. _Psychological Review_, _84_, 216-230.

Cronbach, L. J. (1975). Beyond the two disciplines of scientific psychology. _American Psychologist_, _30_, 116-127.

Davis, M. S. (1971). That's interesting: Towards a phenomenonology of sociology and a sociology of phenomenonology. Philosophy of Social Science, 1, 309-344.

Einhorn, H. J., & Hogarth, R. M. (1978). Confidence in judgment: Persistence of the illusion of validity. Psychological Review, 85, 395-416.

Einhorn, H. J., & Hogarth, R. M. (1981). Behavioral decision theory: Processes of judgment and choice. Annual Review of Psychology, 32, 53-88.

Einhorn, H. J., & Hogarth, R. M. (1982a). Prediction, diagnosis, and causal thinking in forecasting. Journal of Forecasting, 1, 23-36.

Einhorn, H. J., & Hogarth, R. M. (1982b). Reply to commentaries on: Behavioral decision theory. In G. R. Ungson & D. N. Braunstein (Eds.), Decision making: An interdisciplinary enquiry. Boston, MA: Kent.

Einhorn, H. J., & Hogarth, R. M. (1985). Probable cause: A decision making framework. University of Chicago, Center for Decision Research.

Einhorn, H. J., Kleinmuntz, D. N., & Kleinmuntz, B. (1979). Linear regression and process-tracing models of judgment. Psychological Review, 86, 465-485.

Einstein, A., & Infeld, L. (1938). The evolution of physics. New York: Simon & Schuster.

Fama, E. F. (1970). Efficient capital markets: A review of theory and empirical work. Journal of Finance, 35, 383-417.

Festinger, L., Riecken, H. W., Jr., & Schachter, S. (1956). When prophecy fails. Minneapolis: University of Minnesota Press.

Friedman, M. (1956). The methodology of positive economics. In Essays in positive economics. Chicago, IL: University of Chicago Press.

Guttman, L. (1982). Lecture given at Department of Statistics, University of Chicago.

Hammond, K. R. (1978). Psychology's scientific revolution: Is it in danger? (Report No. 211.) University of Colorado, Center for Research on Judgment and Policy.

Hofstee, W. K. B. (1984). Methodological decision rules as research policies: A betting reconstruction of empirical research. Acta Psychologica, 56, 93-109.

Hogarth, R. M. (1974). Monozygotic and dizygotic twins reared together: Sensitivity of heritability estimates. The British Journal of Mathematical and Statistical Psychology, 27, 1-13.

Hogarth, R. M. (1981). Beyond discrete biases: Functional and dysfunctional aspects of judgmental heuristics. Psychological Bulletin, 90, 197-217.

Hogarth, R. M. (1982). On the surprise and delight of inconsistent responses. In R. M. Hogarth (Ed.), Question framing and response consistency. San Francisco, CA: Jossey-Bass.

Hogarth, R. M., & Kunreuther, H. (1985). Ambiguity and insurance decisions. American Economic Review (AEA Papers and Proceedings), 75(2), 386-390.

Kahneman, D., & Tversky, A. (1979). Prospect theory: An analysis of decision under risk. Econometrica, 47, 263-291.

Kershaw, D. N. (1975). The New Jersey negative income tax experiment: A summary of the design, operations and results of the first large-scale social science experiment. In G. M. Lyons (Ed.), Social research and public policies. Hanover, NH: Dartmouth College.

Klayman, J., & Ha, Y.-W. (1985). Confirmation, disconfirmation, and information in hypothesis testing. University of Chicago, Center for Decision Research.

Kunreuther, H. et al. (1978). Disaster insurance protection: Public policy lessons. New York: Wiley.

Kuttner, R. (1985, February). The poverty of economics. The Atlantic Monthly, 74-84.

Lombardo, M. P. (1985). Mutual restraint in tree swallows: A test of the TIT FOR TAT model of reciprocity. Science, 227, 1363-1365.

Mackie, J. L. (1974). The cement of the universe: A study of causation. Oxford: Oxford University Press.

Mahoney, M. J. (1977). Publication prejudices: An experimental study of confirmatory biases in the peer review system. Cognitive Therapy and Research, 1, 161-175.

Maruyama, M. (1963). The second cybernetics: Deviation-amplifying mutual causal processes. American Scientist, 51, 164-179.

Meehl, P. E. (1978). Theoretical risks and tabular asterisks: Sir Karl, Sir Ronald, and the slow progress of soft psychology. Journal of Consulting and Clinical Psychology, 46, 806-834.

Nisbett, R. E. et al. (1963). The use of statistical heuristics in everyday inductive reasoning. Psychological Review, 90, 339-363.

Nisbett, R. E., & Ross, L. (1980). Human inference: Strategies and shortcomings of social judgment. Englewood Cliffs, NJ: Prentice-Hall.

Payne, J. W. (1982). Contingent decision behavior. Psychological Bulletin, 92, 382-402.

Platt, J. R. (1964). Strong inference. Science, 146, 347-353.

Plott, C. R. (1982). Industrial organization theory and experimental economics. Journal of Economic Literature, 20, 1485-1527.

Serlin, R. C., & Lapsley, D. K. (1985). Rationality in psychological research: The good enough principle. American Psychologist, 40, 73-83.

Simon, H. A. (1979). Rational decision making in business organizations. American Economic Review, 69, 493-513.

Smith, V. L. (1982). Microeconomic systems as an experimental science. American Economic Review, 72, 923-955.

Swieringa, R. J., & Weick, K. E. (1982). An assessment of laboratory experiments in accounting. Journal of Accounting Research, 20 (Supplement), 56-101.

Thaler, R. H. (1980). Toward a positive theory of consumer choice. Journal of Economic Behavior and Organization, 1, 39-60.

Thaler, R. H., & Shefrin, H. M. (1983). Life cycle vs. self-control theories of saving: A look at the evidence. Cornell University.

Thomas, L. (1983). The youngest science: Notes of a medicine watcher. New York: Viking Press.

Tversky, A., & Kahneman, D. (1973). Availability: A heuristic for judging frequency and probability. Cognitive Psychology, 5, 207-232.

Tversky, A., & Kahneman, D. (1974). Judgment under uncertainty: Heuristics and biases. Science, 185, 1124-1131.

Tversky, A., & Kahneman, D. (1983). Extensional versus intuitive reasoning: The conjunction fallacy in probability judgment. Psychological Review, 90, 293-315.

von Neumann, J., & Morgenstern, O. (1947). Theory of games and economic behavior (2nd edition). Princeton, NJ: Princeton University Press.

Wason, P. C., & Johnson-Laird, P. N. (1972). Psychology of reasoning: Structure and content. London: Batsford.

Footnotes

This work was supported by a contract from the Office of Naval Research.

OFFICE OF NAVAL RESEARCH

Engineering Psychology Program

TECHNICAL REPORTS DISTRIBUTION LIST

OSD

CAPT Paul R. Chatelier
Office of the Deputy Under Secretary
of Defense
OUSDRE (E&LS)
Pentagon, Room 3D129
Washington, D. C. 20301

Department of the Navy

Engineering Psychology Program
Office of Naval Research
Code 442EP
800 North Quincy Street
Arlington, VA 22217-5000 (3 copies)

Aviation & Aerospace Technology
Programs
Code 210
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

CDR. Paul E. Girard
Code 252
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Physiology Neurobiology Program
Office of Naval Research
Code 441NP
800 North Quincy Street
Arlington, VA 22217-5000

CAPT. P. M. Curran
Code 270
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Mathematics Program
Code 411MA
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Douglas DePriest
Code 411
Office of Naval Research
800 N. Quincy Street
Arlington, VA 22217-5000

Information Sciences Division
Code 433
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

CDR Kent S. Hull
Helicopter/VTOL Human Factors
Office
MS 239-21
NASA/Ames Research Center
Moffett Field, CA 94035

Dr. Carl E. England
Naval Health Research Center
Environmental Physiology
P. O. Box 85122
San Diego, CA 92138

CAPT William M. Houk
Commanding Officer
Naval Medical R&D Command
Bethesda, MD 20814-5055

Department of the Navy

Special Assistant for Marine
Corps Matters
Code 100M
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Mr. R. Lawson
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

CDR James Offutt
Office of the Secretary of Defense
Strategic Defense Initiative Organization
Washington, D.C. 20301-7100

Director
Technical Information Division
Code 2627
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Melich
Communications Sciences Division
Code 7500
Naval Research Laboratory
Washington, D.C. 23075-5000

Dr. J. S. Lawson, Jr.
4773-C Kahala Avenue
Honolulu, HI 96816

Dr. Neil McAlister
Office of Chief of Naval Operations
Command and Control
OP-094H
Washington, D. C. 20350

Naval Training Equipment Center
ATTN: Technical Library
Orlando, FL 32813

Dr. A. F. Norcio
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, DC 20375-5000

Office of the Chief of Naval
Operations, OP987H
Personnel Logistics Plans
Washington, D. C. 20350

Mr. John Davis
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Department
Code N-71
Naval Training Equipment Center
Orlando, FL 32813

Mr. Norm Beck
Combat Control Systems Department
Code 35
Naval Underwater Systems Center
Newport, RI 02840

Human Factors Engineering
Code 441
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Gary Poock
Operations Research Department
Naval Postgraduate School
Monterey, CA 93940

Dean of Research Administration
Naval Postgraduate School
Monterey, CA 93940

Mr. H. Talkington
Engineering & Computer Science
Code 09
Naval Ocean Systems Center
San Diego, CA 92152

Department of the Navy

Mr. Paul Heckman
Naval Ocean Systems Center
San Diego, CA 92152

Dr. Ross Pepper
Naval Ocean Systems Center
Hawaii Laboratory
P. O. Box 997
Kailua, HI 96734

Dr. A. L. Slafkosky
Scientific Advisor
Commandant of the Marine Corps
Code RD-1
Washington, D. C. 20380

Dr. L. Chmura
Computer Sciences & Systems
Code 7592
Naval Research Laboratory
Washington, D.C. 20375-5000

Dr. Michael Letsky
Office of the Chief of Naval
Operations (OP-01B7)
Washington, D.C. 20350

Professor Douglas E. Hunter
Defense Intelligence College
Washington, D.C. 20374

CDR C. Hutchins
Code 55
Naval Postgraduate School
Monterey, CA 93940

Human Factors Technology Administration
Office of Naval Technology
Code MAT 0722
800 North Quincy Street
Arlington, VA 22217-5000

CDR Tom Jones
Naval Air Systems Command
Human Factors Programs
NAVAIR 330J
Washington, D.C. 20361

Commander
Naval Air Systems Command
Crew Station Design
NAVAIR 5313
Washington, D. C. 20361

Mr. Philip Andrews
Naval Sea Systems Command
NAVSEA 61R
Washington, D. C. 20362

Commander
Naval Electronics Systems Command
Human Factors Engineering Branch
Code 81323
Washington, D. C. 20360

Aircrew Systems Branch
Systems Engineering Test
Directorate
U.S. Naval Test Center
Patuxent River, MD 20670

Mr. Milon Essoglou
Naval Facilities Engineering
Command
R&D Plans and Programs
Code 03T
Hoffman Building II
Alexandria, VA 22332

CAPT Robert Biersner
Naval Biodynamics Laboratory
Michoud Station
Box 29407
New Orleans, LA 70189

Dr. Arthur Bachrach
Behavioral Sciences Department
Naval Medical Research Institute
Bethesda, MD

Dr. George Moeller
Human Factors Engineering Branch
Naval Submarine Base
Submarine Medical Research Lab.
Groton, CT 06340

Department of the Navy

Head
Aerospace Psychology Department
Naval Aerospace Medical Research Lab
Pensacola, FL 32508

Commanding Officer
Naval Health Research Center
San Diego, CA 92152

Dr. Jerry Tobias
Auditory Research Branch
Submarine Medical Research Lab
Naval Submarine Base
Groton, CT 06340

Dr. Robert Blanchard
Code 17
Navy Personnel Research and
Development Center
San Diego, CA 92152-6800

LCDR T. Singer
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

Mr. Stephen Merriman
Human Factors Engineering Division
Naval Air Development Center
Warminster, PA 18974

LT. Dennis McBride
Human Factors Branch
Pacific Missle Test Center
Point Mugu, CA 93042

Dr. Kenneth L. Davis
Code 414
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

LCDR R. Carter
Office of Chief on Naval Operations
(OP-01B)
Washington, D.C. 20350

Dean of the Academic Departments
U. S. Naval Academy
Annapolis, MD 21402

CDR W. Moroney
Naval Air Development Center
Code 602
Warminster, PA 18974

Human Factor Engineering Branch
Naval Ship Research and Development
Center, Annapolis Division
Annapolis, MD 21402

Dr. Harry Crisp
Code N 51
Combat Systems Department
Naval Surface Weapons Center
Dahlgren, VA 22448

Mr. John Quirk
Naval Coastal Systems Laboratory
Code 712
Panama City, FL 32401

Human Factors Branch
Code 3152
Naval Weapons Center
China Lake, CA 93555

Dr. Charles Holland
Office of Naval Research Branch Office
London
Box 39
EPO New York 09510

Dr. Rabinder N. Madan
Code 414
Office of Naval Research
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Eugene E. Gloye
ONR Detachment
1030 East Green Street
Pasadena, CA 91106-2485

Dr. David Mizell
ONR Detachment
1030 Green Street
Pasadena, CA 91106-2485

Dr. Glen Allgaier
Artificial Intelligence Branch
Code 444
Naval Electronics Ocean System Center
San Diego, CA 92152

Dr. Steve Sacks
Naval Electronics Systems Command
Code 61R
Washington, D.C. 20363-5100

Dr. Sherman Gee
Command and Control Technology,(MAT 0721)
Office of Naval Technology,
800 North Quincy Street
Arlington, VA 22217-5000

Dr. Robert A. Fleming
Human Factors Support Group
Naval Personnel Research & Development Ctr.
1411 South Fern Street
Arlington, VA 22202

Dr. Dick Kelly
Human Factors Division, Code 17
Naval Personnel Research & Development Center
San Diego, CA 92152-6800

Department of the Army

Dr. Edgar M. Johnson
Technical Director
U.S. Army Research Institute
Alexandria, VA 22333-5600

Technical Director
U.S. Army Human Engineering Laboratory
Aberdeen Proving Ground, MD 21005

Director, Organizations and Systems Research Laboratory
U.S. Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Dr. Robert M. Sasmor
Director, Basic Research
Army Research Institute
5001 Eisenhower Avenue
Alexandria, VA 22333-5600

Department of the Air Force

Dr. Kenneth R. Boff
AF AMRL/HE
Wright-Patterson AFB, OH 45433

Dr. A. Fregly
U. S. Air Force Office of Scientific Research
Life Science Directorate, NL
Bolling Air Force Base
Washington, D.C. 20332-6448

Mr. Charles Bates, Director
Human Engineering Division
USAF AMRL/HES
Wright-Patterson AFB, OH 45433

Dr. Earl Alluisi
Chief Scientist
AFHRL/CCN
Brooks Air Force Base, TX 78235

Dr. R. K. Dismukes
Associate Director for Life Sciences
AFSOR
Bolling AFB
Washington, D.C. 20032-6448

Foreign Addresses

Dr. A. D. Baddeley
Director, Applied Psychology Unit
Medical Research Council
15 Chaucer Road
Cambridge, CB2 2EF England

Dr. Kenneth Gardner
Applied Psychology Unit
Admiralty Marine Tech. Estab.
Teddington, Middlesex
TW11 OLN
England

Other Government Agencies

Dr. M. C. Montemerlo
Information Sciences &
Human Factors Code RC
NASA HQS
Washington, D.C. 20546

Dr. Alan Leshner
Deputy Division Director
Division of Behavioral and
Neural Sciences
National Science Foundation
1800 G. Street, N.W.
Washington, D.C. 20550

Defense Technical Information
Center
Cameron Station, Bldg. 5
Alexandria, VA 22314 (12 copies)

Dr. Clinton Kelly
Defense Advanced Research
Projects Agency
1400 Wilson Blvd.
Arlington, VA 22209

Other Organizations

Dr. Harry Snyder
Dept. of Industrial Engineering
Virginia Polytechnic Institute
and State University
Blacksburg, VA 24061

Dr. Amos Tversky
Dept. of Psychology
Stanford University
Stanford, CA 94305

Dr. Amos Freedy
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Jesse Orlansky
Institute for Defense Analyses
1801 N. Beauregard Street
Alexandria, VA 22311

Dr. J. O. Chinnis, Jr.
Decision Science Consortium,
Inc.
7700 Leesburg Pike
Suite 421
Falls Church, VA 22043

Dr. T. B. Sheridan
Dept. of Mechanical
Engineering
Massachusetts Institute of
Technology
Cambridge, MA 02139

Dr. Daniel Kahneman
The University of British
Department of Psychology
#154-2053 Main Mall
Vancouver, British Columbia
Canada V6T 1Y7

Dr. Stanley Deutsch
NAS-National Research Council
(COHF)
2101 Constitution Avenue, N.W.
Washington, D.C. 20418

Dr. Meredith P. Crawford
American Psychological
Association
Office of Educational Affairs
1200 17th Street N.W.
Washington, D.C. 20036

Dr. Deborah Boehm-Davis
Department of Psychology
George Mason University
4400 University Drive
Fairfax, VA 22030

Dr. Paul E. Lehner
PAR Technology Corp.
7926 Jones Branch Drive
Suite 170
McLean, VA 22102

Other Organizations

Dr. Babur M. Pulat
Department of Industrial Engineering
North Carolina A&T State University
Greensboro, NC 27411

Dr. Stanley N. Roscoe
University of Colorado
Boulder, CO 80309

Dr. James H. Howard, Jr.
Department of Psychology
Catholic University
Washington, D. C. 20064

Dr. William Howell
Department of Psychology
Rice University
Houston, TX 77001

Dr. Christopher Wickens
Department of Psychology
University of Illinois
Urbana, IL 61801

Dr. Robert Wherry
Analytics, Inc.
2500 Maryland Road
Willow Grove, PA 19090

Dr. Edward R. Jones
Chief, Human Factors Engineering
McDonnell-Douglas Astronautics Co.
St. Louis Division
Box 516
St. Louis, MO 63166

Dr. Lola L. Lopes
Department of Psychology
University of Wisconsin
Madison, WI 53706

Dr. Stanley N. Roscoe
New Mexico State University
Box 5095
Las Cruces, NM 88003

Mr. Joseph G. Whol
Alphatech, Inc.
3 New England Executive Park
Burlington, MA 10803

Dr. Marvin Cohen
Decision Science Consortium, Inc.
Suite 721
7700 Leesburg Pike
Falls Church, VA 22043

Dr. William R. Utal
NOSC-Hawaii
Box 997
Kailua, HI 96734

Dr. William B. Rouse
School of Industrial and Systems
Engineering
Georgia Institute of Technology
Atlanta, GA 30332

Ms. Denise Benel
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Dr. Andrew P. Sage
Assoc. V. P. for Academic Affairs
George Mason University
4400 University Drive
Fairfax, VA 22030

Dr. Michael Athans
Massachusetts Inst. of Technology
Lab Information & Decision Systems
Cambridge, MA 01803

Other Organizations

Dr. Richard Pew
Bolt Beranek & Newman, Inc.
50 Moulton Street
Cambridge, MA 02238

Dr. Leonard Adelman
PAR Technology Corp.
Suite 170
7962 Jones Branch Drive
McLean, VA 22102

Dr. Douglas Towne
University of Southern California
Behavioral Technology Lab
1845 South Elena Avenue, Fourth Floor
Redondo Beach, CA 90277

Dr. James P. Jenkins
Nuclear Regulatory Commission
Washington, D.C. 20555

Dr. John Payne
Graduate School of Business
Administration
Duke University
Durham, NC 27706

Dr. Charles Gettys
Department of Psychology
University of Oklahoma
455 West Lindsey
Norman, OK 73069

Dr. Azad Madni
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91364

Dr. Tomaso Poggio
Massachusetts Institute of Tech.
Center for Biological Information
Processing
Cambridge, MA 02139

Dr. Baruch Fischhoff
Perceptronics, Inc.
6271 Variel Avenue
Woodland Hills, CA 91367

Dr. Robert A. Hummel
New York University
Courant Inst. of Mathematical
Sciences
251 Mercer Street
New York, New York 10012

Dr. H. McI. Parsons
Essex Corporation
333 N. Fairfax Street
Alexandria, VA 22314

Dr. Paul Solvic
Decision Research
1201 Oak Street
Eugene, OR 97401

Dr. David Castanon
ALPHATECH, Inc.
111 Middlesex Turnpike
Burlington, MA 01803

Dr. E. Douglas Jensen
Carnegie-Mellon University
Computer Science Dept.
Pittsburgh, PA 15213

Dr. David Noble
Engineering Research Assoc.
8616 Westwood Center Drive
McLean, VA 22180

Dr. Bruce Hamill
The Johns Hopkins Univ.
Applied Physics Lab
Laurel, MD 20707

Dr. A. Ephremides
University of Maryland
Electrical Engineering Dept.
College Park, MD 20742

END

FILMED

12-85

DTIC